

Notice of Allowability**Application No.**

10/501,397

Applicant(s)

LINDH ET AL.

Examiner

BELINDA XUE

Art Unit

2163

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to Amendment filed on 2/25/2009.
2. ☒ The allowed claim(s) is/are 1,2,4-10,12-14,17-20,27 and 28.
3. ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some* c) ☐ None of the:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. ____.
3. ☒ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

* Certified copies not received: ____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.

THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

4. ☐ A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.
5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
- (a) ☐ including changes required by the Notice of Draftsperson's Patent Drawing Review (PTO-948) attached
- 1) ☐ hereto or 2) ☐ to Paper No./Mail Date ____.
- (b) ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date ____.
- Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).
6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3. ☐ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date ____
4. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
5. ☐ Notice of Informal Patent Application
6. ☒ Interview Summary (PTO-413), Paper No./Mail Date ____
7. ☒ Examiner's Amendment/Comment
8. ☒ Examiner's Statement of Reasons for Allowance
9. ☐ Other ____.

/Belinda Xue/
Examiner, Art Unit 2163

/don wong/
Supervisory Patent Examiner, Art Unit 2163

DETAILED ACTION

Examiner's Amendment

1. An examiner's amendment to the record appears below. Should the changes and/or additions be unacceptable to applicant, an amendment may be filed as provided by 37 CFR 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.
2. Authorization for this examiner's amendment was given in a telephone interview with Michael A. Schwartz, Reg. No. 40,161 on 5/8/2009.

In the claims:

Please replace claims 1-2, 4-10, 12-14, 17-18, 20 and 27-28 with amended claims 1-2, 4-10, 12-14, 17-18, 20 and 27-28 respectively.

Art Unit: 2163

1. (currently amended) A method of processing digitized textual information in a computerized database system, the information being organized in terms, documents and document corpora, where each document contains at least one term and each document corpus contains at least one document, the method comprising:

generating, by using a computer, a concept vector for each document in a document corpus wherein the concept vector conceptually classifying the contents of the document on a relatively compact format,

generating, for each term in the document corpus, a term-to-concept vector describing a relationship between the term and each of the concept vectors wherein the term-to-concept vectors being generated on basis of the concept vectors, comprises:

receiving the term-to-concept vectors for the document corpus and on basis thereof generating a term-term matrix describing a term-to-term relationship between the terms in the document corpus, [[and]] wherein the generation of the term-term matrix comprises: retrieving, for each term in each combination of two unique terms in the document corpus, a respective term-to-concept vector, generating a relation vector describing the relationship between the terms in the each combination of two unique terms, each component in the relation vector being equal to a lowest component value of corresponding component values in the term-to-concept vectors, generating a relationship value for the each combination of two unique terms as the sum of all component values in the corresponding relation vector, and generating a matrix containing the

Art Unit: 2163

relationship values of all combinations of two unique terms in the document corpus,

processing the term-term matrix into processed textual information and displaying the processed textual information via a user output interface[.], and

displaying the processed textual information as a distance graph in which each term constitutes a node wherein the node representing a first term is connected to one or more other nodes representing secondary terms to which the first term has a conceptual relationship of at least a specific strength, and a relevance measure between the first term and at least one second term is represented by a minimum number of node hops between the first term and the at least one second term.

2. (currently amended) [[A]] The method according to claim 1, wherein each document in the document corpus being associated with a document-concept matrix representing at least one concept element whose relevance with respect to the document is described by a weight factor, the generation of each term-to-concept vector comprises:

identifying a term-relevant set of documents in the document corpus, each document in the term-relevant set containing at least one occurrence of the term,

calculating a term weight for the term in each of the documents in the term-relevant set,

Art Unit: 2163

retrieving a respective concept vector being associated with each document in the term-relevant set where the term weight exceeds a first threshold value,

selecting a relevant set of concept vectors including any concept vector in which at least one concept component exceeds a second threshold value,

calculating a non-normalized term-to-concept vector as the sum of all concept vectors in the relevant set, and

normalizing the non-normalized term-to-concept vector.

4. (currently amended) [[A]] The method according to claim 1 wherein the method further comprises the steps of:

calculating a statistical co-occurrence value between the each combination of two unique terms in the document corpus, the statistical co-occurrence value describing a dependent probability that a certain second term exists in a document provided that a certain first term exists in the document, and incorporating the statistical co-occurrence values into the term-term matrix to represent lexical relationships between the terms in the document corpus.

5. (currently amended) [[A]] The method according to claim 1 wherein the method further comprises the step of:

displaying the processed textual information on a format being adapted for human comprehension.

Art Unit: 2163

6. (currently amended) [[A]] The method according to claim 5, wherein the displaying step further comprises involving presentation of:

at least one document identifier specifying a document being relevant with respect to at least one term in a query,

at least one term being related to a term in [[a]] the query, and

a conceptual distribution representing a conceptual relationship between two or more terms in the document corpus, the conceptual distribution being based on shared concepts which are common to said terms.

7. (currently amended) [[A]] The method according to claim 5 wherein the displaying step further comprises involving presentation of at least one document identifier specifying a document being relevant with respect to at least one term in a query in combination with at least one user specified concept.

8. (currently amended) [[A]] The method according to claim 6 wherein the method further comprises the step of:

selecting the at least one user specified concept from the shared concepts in the conceptual distribution.

9. (currently amended) [[A]] The method according to claim 5 wherein the method further comprises the step of:

illustrating the conceptual relationship between [[a]] the first term and the at least one second term by means of a respective relevance measure being associated with the at least one second term in respect of the first term.

10. (currently amended) [[A]] The method according to claim 9, wherein the method further comprises the step of:

displaying the processed textual information on a graphical format which visualizes the strength in the conceptual relationship between at least two terms.

12. (currently amended) [[A]] The method according to claim 9 wherein the method further comprises the step of:

displaying the processed textual information as a distance graph in which each term constitutes a node wherein the node representing [[a]] the first term is connected to one or more other nodes representing secondary terms to which the first term has a conceptual relationship, each connection is associated with an edge weight representing the strength of a conceptual relationship between the first term and a particular secondary term, and the relevance measure between the first term and [[a]] the particular secondary term is represented by an accumulation of the edge weights being associated with the connections constituting a minimum number node hops between the first term and the particular secondary term.

13. (currently amended) [[A]] The method according to claim 1, wherein each term further comprises:

- a single word,
- a proper name,
- a phrase, and
- a compound of single words.

14. (currently amended) [[A]] The method according to claim 1 further comprises the step of updating the document corpus with added data in form of at least one new document by means of

- identifying any added terms in the new document which lack a representation in the document corpus,

- identifying any existing terms in the new document which were represented in the document corpus before adding the at least one new document,

- retrieving, for each of the existing terms, a corresponding concept vector, generating a new concept vector with respect to the at least one new document as a sum of the corresponding concept vectors,

- normalizing the new concept vector into a normalized new concept vector, and

- assigning the normalized new concept vector to each of the added terms in the new document.

Art Unit: 2163

17. (currently amended) A computer-implemented search engine, embedded on a computer readable storage medium, for processing an amount of digitized textual information and extracting data there from, the information being organized in terms, documents and document corpora, where each document contains at least one term and each document corpus contains at least one document, comprising:

an interface configured to receive a query from a user, and

a processing unit configured to process a document corpus on basis of the query and return processed textual information being relevant to the query said process involving

generating a concept vector for each document in the document corpus, the concept vector conceptually classifying ~~the~~ contents of the document on a relatively[[']] compact format, and

generating, for each term in the document corpus, a term-to-concept vector describing a relationship between the term and each of the concept vectors, wherein the processing unit in turn comprises:

a processing module configured to receive the term-to-concept vectors for the document corpus and on basis thereof generate a term-term matrix describing a term-to-term relation-ship between the terms in the document corpus, [[and]] wherein the generation of the term-term matrix comprises: retrieving, for each term in each combination of two unique terms in the document corpus, a respective term-to-concept vector, generating a relation vector describing the relationship between the terms in the each combination of

two unique terms, each component in the relation vector being equal to a lowest component value of corresponding component values in the term-to-concept vectors, generating a relationship value for the each combination of two unique terms as the sum of all component values in the corresponding relation vector, and generating a matrix containing the relationship values of all combinations of two unique terms in the document corpus,

an exploring module configured to receive the query and the term-term matrix, and on basis of the query process the term-term matrix into the processed textual information[[.]] , and

a display module configured to display the processed textual information as a distance graph in which each term constitutes a node wherein the node representing a first term is connected to one or more other nodes representing secondary terms to which the first term has a conceptual relationship of at least a specific strength, and a relevance measure between the first term and at least one second term is represented by a minimum number of node hops between the first term and the at least one second term,

18. (currently amended) A computer-implemented database system comprising:
a processor;

memory holding an amount of digitized textual information being organized in terms, documents and document corpora, wherein each document contains at least one term and each document corpus contains at least one document, wherein each document in a document corpus being associated with

Art Unit: 2163

concept vector which conceptually classifies the contents of the document on a relatively compact format, and wherein each term in the document corpus being associated with a term-to-concept vector describing a relationship between the term and each of the concept vectors, ~~wherein it is configured to deliver~~ delivering the term-to-concept vectors to a search engine for processing an amount of digitized textual information and extracting data there from, the information being organized in terms, documents and document corpora, where each document contains at least one term and each document corpus contains at least one document, and

computer program instructions implementing: comprising:-

an interface configured to receive a query from a user, and

a processing unit configured to process a document corpus on basis of the query and return processed textual information being relevant to the query said process involving

generating a concept vector for each document in the document corpus, the concept vector conceptually classifying the contents of the document on a relatively[["]] compact format, and

generating, for each term in the document corpus, a term-to-concept vector describing a relationship between the term and each of the concept vectors, wherein the processing unit in turn comprises:

a processing module configured to receive the term-to-concept vectors for the document corpus and on basis thereof generate a term-term matrix describing a term-to-term relation-ship between the terms in the document

Art Unit: 2163

corpus, [[and]] wherein the generation of the term-term matrix comprises:
retrieving, for each term in each combination of two unique terms in the
document corpus, a respective term-to-concept vector, generating a relation
vector describing the relationship between the terms in the each combination of
two unique terms, each component in the relation vector being equal to a lowest
component value of corresponding component values in the term-to-concept
vectors, generating a relationship value for the each combination of two unique
terms as the sum of all component values in the corresponding relation vector,
and generating a matrix containing the relationship values of all combinations of
two unique terms in the document corpus,

an exploring module configured to receive the query and the term-term matrix, and on basis of the query process the term-term matrix into the processed textual information^{[[.]]}, and

a display module configured to display the processed textual information
as a distance graph in which each term constitutes a node wherein the node
representing a first term is connected to one or more other nodes representing
secondary terms to which the first term has a conceptual relationship of at least a
specific strength, and a relevance measure between the first term and at least
one second term is represented by a minimum number of node hops between
the first term and the at least one second term.

19. (currently amended) [[A]] The computer-implemented database system according to claim 18 further comprising an iterative term-to-concept engine

Art Unit: 2163

configured to receive fresh digitized textual information added to the database and on basis of this information

generate concept vectors for any added document, and

generate a term-to-concept vector describing a relationship between any added term and each of the concept vectors.

20. (currently amended) A server computer system for providing data processing services in respect of digitized textual information, wherein the server comprises:

a processor;

memory for storing computer program instructions and data; and

computer program instructions stored in the memory for implementing;

a search engine for processing an amount of digitized textual information and extracting data there from, the information being organized in terms, documents and document corpora, where each document contains at least one term and each document corpus contains at least one document, comprising an interface configured to receive a query from a user, and a processing unit configured to process a document corpus on basis of the query and return processed textual information being relevant to the query said process involving generating a concept vector for each document in the document corpus, the concept vector conceptually classifying the contents of the document on a relatively compact format, and generating, for each term in the document corpus, a term-to-concept vector describing a relationship between the term and each of the concept vectors, wherein the processing unit in turn comprises a processing

Art Unit: 2163

module configured to receive the term-to-concept vectors for the document corpus and on basis thereof generate a term-term matrix describing a term-to-term relation-ship between the terms in the document corpus, [[and]] wherein the generation of the term-term matrix comprises: retrieving, for each term in each combination of two unique terms in the document corpus, a respective term-to-concept vector, generating a relation vector describing the relationship between the terms in the each combination of two unique terms, each component in the relation vector being equal to a lowest component value of corresponding component values in the term-to-concept vectors, generating a relationship value for the each combination of two unique terms as the sum of all component values in the corresponding relation vector, and generating a matrix containing the relationship values of all combinations of two unique terms in the document corpus, an exploring module configured to receive the query and the term-term matrix, and on basis of the query process the term-term matrix into the processed textual information, [[and]] a display module configured to display the processed textual information as a distance graph in which each term constitutes a node wherein the node representing a first term is connected to one or more other nodes representing secondary terms to which the first term has a conceptual relationship of at least a specific strength, and a relevance measure between the first term and at least one second term is represented by a minimum number of node hops between the first term and the at least one second term, and

a communication interface towards a database system ~~according to claim~~
~~48- holding an amount of digitized textual information and configured to deliver~~
~~the term-to concept vectors to the search engine.~~

27. (currently amended) A computer system comprising: a processor for
executing computer program instructions, a memory for storing computer
program instructions and computer program instructions comprising software for
~~performing a method of~~ processing digitized textual information, the information
being organized in terms, documents and document corpora, where each
document contains at least one term and each document corpus contains at least
one document, the digitized textual information processed by the method
~~comprising:~~

generating a concept vector for each document in a document corpus
wherein the concept vector conceptually classifying the contents of the document
on a relatively compact format,

generating, for each term in the document corpus, a term-to-concept
vector describing a relationship between the term and each of the concept
vectors wherein the term-to-concept vectors being generated on basis of the
concept vectors,

receiving the term-to-concept vectors for the document corpus and on
basis thereof generating a term-term matrix describing a term-to-term
relationship between the terms in the document corpus, ~~[[and]]~~ wherein the
generation of the term-term matrix comprises: retrieving, for each term in each

Art Unit: 2163

combination of two unique terms in the document corpus, a respective term-to-concept vector, generating a relation vector describing the relationship between the terms in the each combination of two unique terms, each component in the relation vector being equal to a lowest component value of corresponding component values in the term-to-concept vectors, generating a relationship value for the each combination of two unique terms as the sum of all component values in the corresponding relation vector, and generating a matrix containing the relationship values of all combinations of two unique terms in the document corpus,

processing the term-term matrix into processed textual information and displaying the processed textual information via a user output interface[.], and displaying the processed textual information as a distance graph in which each term constitutes a node wherein the node representing a first term is connected to one or more other nodes representing secondary terms to which the first term has a conceptual relationship of at least a specific strength, and a relevance measure between the first term and at least one second term is represented by a minimum number of node hops between the first term and the at least one second term.

28. (currently amended) A computer program product ~~comprising a~~ stored in a computer readable storage medium, having the computer program product comprising:

computer program instructions recorded thereon for causing a computer to ~~perform a method of processing process~~ digitized textual information, the information being organized in terms, documents and document corpora, where each document contains at least one term and each document corpus contains at least one document, ~~method comprising the digitized textual information~~ processed by:

generating a concept vector for each document in a document corpus wherein the concept vector conceptually classifying the contents of the document on a relatively compact format,

generating, for each term in the document corpus, a term-to-concept vector describing a relationship between the term and each of the concept vectors wherein the term-to-concept vectors being generated on basis of the concept vectors,

receiving the term-to-concept vectors for the document corpus and on basis thereof generating a term-term matrix describing a term-to-term relationship between the terms in the document corpus, ~~[[and]]~~ wherein the generation of the term-term matrix comprises: retrieving, for each term in each combination of two unique terms in the document corpus, a respective term-to-concept vector, generating a relation vector describing the relationship between the terms in the each combination of two unique terms, each component in the relation vector being equal to a lowest component value of corresponding component values in the term-to-concept vectors, generating a relationship value for the each combination of two unique terms as the sum of all component values

Art Unit: 2163

in the corresponding relation vector, and generating a matrix containing the relationship values of all combinations of two unique terms in the document corpus,

processing the term-term matrix into processed textual information and displaying the processed textual information via a user output interface[.]. and displaying the processed textual information as a distance graph in which each term constitutes a node wherein the node representing a first term is connected to one or more other nodes representing secondary terms to which the first term has a conceptual relationship of at least a specific strength, and a relevance measure between the first term and at least one second term is represented by a minimum number of node hops between the first term and the at least one second term.

Reasons For Allowance

3. Claims 1-2, 4-10, 12-14, 17-18, 20 and 27-28 are allowed.

The prior art of record, alone or in combination, does not teach or fairly suggest the combination of the steps as recited in independent claims 1, 17, 18, 20, 27 and 28, wherein "receiving the term-to-concept vectors for the document corpus and on basis thereof generating a term-term matrix describing a term-to-term relationship between the terms in the document corpus, wherein the generation of the term-term matrix comprises: retrieving, for each term in each combination of two unique terms in the document corpus, a respective term-to-concept vector,

Art Unit: 2163

generating a relation vector describing the relationship between the terms in the each combination of two unique terms, each component in the relation vector being equal to a lowest component value of corresponding component values in the term-to-concept vectors, generating a relationship value for the each combination of two unique terms as the sum of all component values in the corresponding relation vector, and generating a matrix containing the relationship values of all combinations of two unique terms in the document corpus, and displaying the processed textual information as a distance graph in which each term constitutes a node wherein the node representing a first term is connected to one or more other nodes representing secondary terms to which the first term has a conceptual relationship of at least a specific strength, and a relevance measure between the first term and at least one second term is represented by a minimum number of node hops between the first term and the at least one second term."

The dependent claims bring definite, further limiting, and fully enable by the specification are also allowed.

4. Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

Contact Information

5. Any inquiry concerning this communication or earlier communications from the examiner should be directed to BELINDA XUE whose telephone number is (571)270-1762. The examiner can normally be reached on Monday-Friday, 9:00am-5:30pm EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Don Wong can be reached on 571-272-1834. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

May 9, 2009

/Belinda Xue/
Examiner,
Art Unit 2163

Art Unit: 2163

/don wong/

Supervisory Patent Examiner, Art Unit 2163